DOCUMENT RESUME

ED 079 747 CS 200 631

AUTHOR Pollio, Howard R.
TITLE Figurative Language: A Neglected Aspect of the Elementary School Language Arts Curriculum. Final Report.
INSTITUTION Tennessee Univ., Knoxville. Dept. of Psychology.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
BUREAU NO BR-1-D-049
PUB DATE May 73
GRANT OEG-4-71-0066
NOTE 41p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Composition (Literary); *Elementary Education; *Figurative Language; Grade 3; Grade 4; Grade 5; *Language Arts; *Language Research; Language Usage; Metaphors; Reading Materials

ABSTRACT

Given the rather sparse and somewhat confusing literature dealing with the topic or metaphor and metaphoric usage, this study attempted to (1) determine the basal level of metaphorical usage in grades 3, 4, and 5; (2) evaluate a series of books ("Making It Strange") designed to promote and augment figurative usage in these grades; (3) determine the differences between groups exposed to this series and those using the normal curriculum of the school; and (4) uncover factors in an individual child's test profile that might predispose him to differential metaphorical usage. Conclusions based on the results of the study included: (1) Children use figures of speech (both novel and frozen) in their compositions as early as the third grade, and novel usage seems to decrease over grade level. (2) The experimental program using "Making It Strange" increased the frequency of occurrence of novel figurative usage in the children exposed to this series of compared to their initial output, and as compared to the control groups. Exposure to this series, however, had no effect on the production of frozen figures of speech. (3) There is a strong relationship between frozen figures and length of composition, thereby suggesting that frozen figures might be considered simply as learned vocabulary items. (HOD)

ED 079747

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Final Report

Project No. 1-D-049

Grant No. OEG-4-71-0066

Howard R. Pollio

Department of Psychology

University of Tennessee

Knoxville, Tennessee 37916

FIGURATIVE LANGUAGE: A NEGLECTED ASPECT OF THE ELEMENTARY SCHOOL LANGUAGE ARTS CURRICULUM

May, 1973

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Office of Education

National Institute of Education

OS 200 631

ABSTRACT

Two classes per grade (one experimental and one control) for Grades 3, 4, and 5 were pretested to determine the developmental level of metaphorical usage prior to experimental treatment. The Making It Strange series was used with experimental classes for a period of twenty days. At the close of this phase a posttest was administered to all classes.

Pretest compositions were examined and instances of figurative language recorded. Reliability measures showed that trained raters could record such usage quite reliably. Analyses of pretest data showed that children produced a greater number of frozen than novel figures, and that production of novel figures decreased over the grades studied. Correlational data showed that frozen figures and composition length were more strongly related than novel figures and composition length. No relationship was found between frozen and novel figures for individual children.

Posttest results showed metaphoric usage increased over all grades, although only increases in novel usage were statistically significant. Experimental groups using the experimental series showed large and significant increases in novel figures, while control groups did not. Correlational data showed a positive relationship between frozen figures and composition length and a weaker relationship between novel figures and composition length. Again, no relationship was found between number of novel and frozen figures.

I. Introduction

Every young adult nearing graduation from high school knows what a metaphor is -- or does he? A student can usually give a vague definition of th term and perhaps supply a few examples from poetry or novels that he has read; ut the question still remains, does the average high school or even college graduate know not only what metaphor is, but what it can do? Does he recognize either the frequency or importance of metaphoric expression in the world around him?

Traditionally, figurative language such as metaphor has been viewed as the province of literature and literary analysis. Metaphors are linguistic devices encountered and dealt with originally in the context of novels or poetry, and only later applied with limited success to the area of written composition. Although this may be a valid approach to figurative language, there is a growing realization that metaphor is something more than simply linguistic ornamentation. So, for example, the poet Elizabeth Sewell (1964) has written:

> (A certain metaphor) though exact will take me no further. I cannot think with it; merely note its exactness and leave it there...It is fitting but not fertile...In great poets for the most part, it is hard to find metaphors of this partial sort. (In great poems) all the figures work, have energy or lend the mind energy to work and to work further. That is to say...they are beautiful, beauty being considered as just such a dynamic heuristic, whether we meet it in the figures of science, those of poetry, or elsewhere. It is exactly such a forward-moving or prophetic energy that the chosen metaphor, within the method in use, has to supply.

As a review of the literature will show, Sewell is not alone in her belief that metaphor should be viewed as a "dynamic heuristic" capable of producing creative problem solving in a wide variety of different contexts; e.g., in fields as diverse as psychotherapy and industrial invention. If figurative language is to be taught in the schools as such a device, it becomes necessary to answer several questions as to the function of metaphor in human thinking, the role of figurative usage in language and speech, the developmental course of metaphoric usage, and the extent to which training in this process will facilitate metaphorical usage in elementary school children.

A. The Functions of Metaphor

Metaphors (and all figurative terms) are often regarded as ornaments in language. One primarily associates the use of metaphor with poetry and, to a lesser degree, with creative fiction. The reason for using figurative language in creative writing is obvious: plain, bare facts written in a plain, cold manner are boring. As Mulder (1959) and Ferguson (1958) both point out, metaphoric expression, and the images it arouses, lend beauty and interest to all writing, whether it be scientific, literary, or otherwise.

Metaphor can also be considered a verbal shorthand enabling us to express thoughts that otherwise could only be expressed by a long and cumbersome phrase or sentence(s) (Stern, 1931). Somewhat along the same line is the observation that metaphors augment vocabulary. Many authorities agree that a large vocabulary is necessary for success in our complex, ever-changing world (Brown, 1959; Upton, 1964a, 1964b; Wiersma, 1959), but even more important is the fact that further vocabulary additions are possible through metaphoric extensions of words already known (Upton, 1964a, 1964b).

Cooper, among others (Armstrong, 1945; Burkland, 1964; Cooper, 1954; Newton, 1964; Upton, 1964a, 1964b), emphasizes that an understanding of figurative language is necessary for critical reading. Given the liberal use of metaphor in

advertising, in politics, and in propaganda; how can we -- or our children -- critically read, evaluate and understand the intentions of a speaker or author if we do not have skill in interpreting figurative language?

Considered from a slightly larger perspective, Stelzer (1965) notes that metaphor often serves as an aid in understanding the interests and ideas of a given historical period or age. He feels that an age can be succinctly characterized by its metaphors and that these, in turn, influence the subject areas to which they are applied. "As a form of comparative analysis, metaphors can structure inquiry, establish relevance and provide an interpretive system." This insight has been cast in the form of a general linguistic hypothesis that has come to be known as Sperber's Law. Basically, Sperber's Law states that an area of intense interest in a culture will become a center for metaphorical expression and extension. "If at a certain time a complex of ideas is so strongly charged with feeling that it causes one word to extend its sphere and change its meaning, we may confidently expect that other words belonging to the same emotional complex will also shift their meaning (as cited in Ullmann, 1963, p. 240)."

Recently, metaphor has been given new emphasis in the area of psychotherapy. The psychoanalist Rudolf Ekstein (1966) feels that interpretation within the metaphor of the patient is often an effective way of dealing with primary process material without lifting it into the context of secondary process, i.e., by keeping it at a distance. He has used this method with some success in dealing with borderline schizophrenic patients. If the relations expressed in the metaphor are interpreted by the therapist within the context of the metaphor itself, it is less threatening to the patient and the communication so vital to therapy can be established and maintained.

Lenrow (1966) shares this view of the role of metaphor in psychotherapy. In addition, he lists seven further functions of metaphor:

1. Metaphors provide a model of willingness to try out novel ways of looking at behavior.
2. Metaphors simplify events in terms of a schema, or concept, that emphasizes some properties more than others.
3. Metaphors give communications an intimate or personal quality because of the concrete referents of metaphorical imagery.
4. Metaphors have a half-playful, half-serious quality that permits the therapist to communicate about intimate characteristics of the patient without appearing as intrusive as a more conventional mode of describing the patient might appear. The dissimilarities between person and metaphoric referents may help the patient to consider the possible similarities without generalized avoidance or defense against new concepts of himself.
5. Metaphors assert the affective equivalence of apparently dissimilar concepts or events. An apt metaphor may permit the patient to observe his own ways of equating situations and thus open possibilities of dealing with the situations as different in important respects.
6. Metaphors highlight subtle social roles that characterize the patient's mode of relating to others when they refer to interactions between an object and its environment.
7. Metaphors transfer readily to new situations that the person enters, or old ones he re-enters. This is because they refer to relational properties rather than to discrete elements, and can thus be applied in a great variety of settings.

Another way in which metaphor, and more specifically, poetry, enter into psychotherapy has been described by J. J. Leedy in his book Poetry Therapy (1969). Leedy believes that communication can be enhanced by giving patients poetry that matches their psychological condition and by encouraging them to write their own poems. Leedy, of course, is using the vocabulary augmenting aspect of metaphor

in psychotherapy. A slightly different but related analysis stressing the therapeutic potential of metaphoric communication in psychotherapy has also been offered by Fine, Pollio, and Simpkinson (1973).

A more radical view of the function of metaphor, however, has been taken by Donald Schon (1963). He states that the process of metaphor, which he calls "the displacement of concepts," is essential to the development of new theories: scientific or otherwise. The displacement of concepts is "the functioning of older theories as metaphors or projective models for new situations." In order to function as a "projective model," older concepts come to be seen in a new light and ultimately change into a newer more usable concept. According to Schon, "there is here the possibility of a new kind of inquiry -- an intellectual history which would consider not the manifest content of theories; but the development of the underlying metaphors; a history of the displacement of theories....It would attempt to describe the patterns of interaction and change in metaphor and theory (p. 192)."

A somewhat similar, though less philosophical and more pragmatic view of metaphor as a heuristic, has been developed by William J. J. Gordon (1961). Gordon, the founder of Synectics, Inc., is the originator of synectic theory and its attendant operational mechanisms. The word "synectics" is a neologism based on the Greek, meaning the "fitting together of diverse elements." Synectic theory deals with the integration of diverse individuals into a "problem-stating, problem-solving" group in an attempt to foster creative invention. It is "an operational theory for the conscious use of the preconscious psychological mechanisms present in man's creative activity." Such research is based on three assumptions:

1. that the creative process in human beings can be concretely described and, further, that sound description should be usable in teaching methodology to increase the creative output of both individuals and groups;
2. that the cultural phenomena of invention in the arts and in science are analogous and are characterized by the same fundamental psychic processes;
3. that individual process in the creative enterprise enjoys a direct analogy in group process.

With these assumptions in mind, Gordon and his co-workers examined the biographies and autobiographies of numerous creative thinkers as well as reports of people in the process of invention. In this manner they isolated what they feel are the significant psychological states present in any creative act. These states are: involvement (closeness and identification with the problem at hand); detachment (the feeling of being removed or cut off); deferment (the sense of putting off a premature attempt at a solution); speculation (the ability to let the mind run free); and autonomy of object (the feeling that comes as the problem nears solution; that the object is autonomous and acting on its own).

Yet knowledge of these states is not enough to enable people to use them in solving new problems. With the help of much in-house research, the Synectics Group has developed an operational procedure for getting problem-solvers to these states. The Synectics process involves making the strange familiar (understanding the problem as given with all of its various ramifications) and making the familiar strange (taking the problem and distorting, inverting, and transposing it, so that it can be viewed from new and different perspectives). There are four mechanisms involved in making the familiar strange with each involving a different metaphoric analysis of the problem:

1. Personal analogy is becoming yourself, one of the objects looked at; and feeling, thinking and acting like that object. It goes beyond mere role-playing in that you can be an inanimate, as well as an animate object or being. This is, essentially, personification.
2. Direct analogy is an actual comparison of parallel facts, knowledge, or technology. In problem-stating, problem-solving situations, analogies from the biological sciences appear to be the most fruitful.

3. A symbolic analogy is a compressed description of the function or elements of the problem as the problem solver views it. It is the poetic response that sums up what has been said in the personal and direct analogy phases. It is, therefore, a direct metaphor.
4. Fantasy analogy is solution of the problem by wish-fulfillment, i.e., wishing the problem solved in any manner whatsoever.

Actual problem solving is done in a group, with different group members entering into the different mechanisms at any and/or all points. According to Gordon, it is possible to recognize when one is nearing problem-solution by the evocation of what he calls the "hedonistic response." This is the very real feeling of pleasure that you get when you are on the right track to solution. He equates it somewhat with what has been described as intuition. Although the whole synectic process sounds like a magical, ritualistic incantation designed to summon the goddess Creativity, its track record has been impressive. When used in industry, improved products and inventions have resulted; when used at Harvard University in connection with science teaching, improved learning seems to have occurred (Gordon, 1965).

B. Metaphor, Language, and Speech

It is quite obvious to any listener (or reader) that metaphor occurs quite often in English. In an attempt to determine if this is true in other languages as well, Asch (1958) investigated the use of metaphorical terms in eight different languages (including English) through an analysis of double-function terms, i.e., terms that refer jointly to psychological and physical properties of people and objects. Basically, Asch was interested in two questions: (1) Do historically independent languages employ the same words to designate physical and psychological properties?; and (2) If so, do languages belonging to different families agree in the detailed pairings they make? The results of this study showed that all of the languages examined did contain terms referring both to physical and psychological qualities. Each language was found to possess some words that referred to the same paired properties as those found in other languages. Not only did all languages considered contain metaphors; they seemed to contain similar metaphors, at least in so far as this particular set of double-function terms was concerned.

Most of us, as native speakers of a language, are often unaware of the fact that we are speaking metaphorically. Consider the following examples: "He is a cold person" (But I'm sure his temperature is the same as yours or mine); "She is bursting with joy" (But where are the fissures on her skin?) and so on. These metaphors are what is known as "dead" or "frozen" metaphors. Dead metaphors are metaphors which have become "frozen" into the ordinary vocabulary of the language. Their meanings are set and are easily understood by any mature speaker of the language (Brown, 1958b; Brown in Bruner, Goodnow and Austin, 1956; Brown, 1965; Stern, 1931; Stewig, 1966). Sometimes, as in the above examples, both the metaphorical and literal meanings of the terms persist; sometimes, however, the original meaning disappears and only the figurative meaning remains. The word "bedes" originally meang "prayers." People would say their "bedes" while counting them out on the small, round balls of the rosary. Gradually, "bedes" or "beads" came to refer to the balls of the rosary. For a while, both senses of the word remained in the lexicon. Now, the meaning of "beads" is so far removed from its original meaning, that we can refer to the round balls strung on a necklace (or on a curtain) as beads (Stern, 1931). This is an example of a semantic change due to a metaphorical relation fostered by contiguity between the constituent concepts. The linguistic history of the English language is replete with examples of this major kind of semantic shift (Brown, 1958b; Brown, 1965; Stern, 1931).

If metaphor has a role in changing meaning, then it also must play a part in the naming of objects. It is, perhaps, easiest to see this role of metaphor

in language if we look at recent additions to our ordinary vocabulary, e.g., telephone - "far-speaker," television - "far-see-er," skyscraper - "scrapes the sky," etc. Brown, 1958a; Stern, 1931). Metaphor clearly plays an important role in extending the vocabulary of a given language.

In terms of more theoretical issues, metaphor has proven to be an extremely difficult topic to handle within the context of semantic analysis. Consider, for example, the now famous class of semantic theory proposed by Katz and Fodor (1963) and extended by Katz (1966) and others (Weinreich, 1966; Johnson, 1970). One major problem, among many others (see Bolinger, 1965), with this class of theory is that it cannot handle metaphoric or figurative language. The markers and distinguishers in Katz and Fodor's dictionary denote specific senses of the word. One must follow one of the paths from the word along the markers to the distinguisher. There are no paths, and no way of making new paths, for metaphoric flexibility or change in meaning.

An attempt has been made by Weinreich (1966) to improve on this particular model. Weinreich postulates a "construal rule" which operates on contradictions, i.e., meanings that do not fit when brought down from the dictionary. As Weinreich put it: "(a construal rule operates by producing) a new semantic entity with a more elaborate structure in which the transferred feature is decisive; but in which the contradictury feature can be accommodated." Take, for example, the sentence "He trues the theorem." Since "trues" is not in the dictionary with the grammatical marker "verb," Katz and Fodor would not allow the sentence and would be forced to call it agrammatical. This is obviously false. The sentence can be meaningfully and consistently interpreted. Weinreich's rule takes the dictionary meaning as a starting point and restates the sentence so as to accomodate the new meaning by changing the form of the sentence to: "He proved the theorem to be true." Although this rule is cumbersome to apply, it does represent an advance over Katz and Fodor's views, yet still falls short of success. Perhaps, future theories of this type will succeed in handling figurative language or, perhaps, all such theories will fail because there is no general nor logical structure to the lexicon. In any case metaphor would seem to hold a key to understanding the organization of vocabulary.

C. The Development of Metaphor

Although much has been written on the development of language in children, there has been only one study directly concerned with the development of figurative usage -- and not an extensive one at that. In this study Asch and Nerlove (1960) used fifty children (five groups with ten subjects in each group), ranging in age from three to twelve. The subjects came from upper middle class homes in the Swarthmore area. Since the experimenters specifically note that they were looking for trends, rather than attempting to provide normative data, they felt that this sample was sufficient for this purpose. Each of the children in this study was interviewed on a one to one basis and questioned about a limited number of double-function terms, i.e., terms such as sweet or hard which refer both to the physical properties of things as well as to the psychological properties of people. The results of this study indicated that mastery of double-function terms followed a regular development course, with young children tending first to use these terms strictly in reference to objects. The psychological sense of a double-function term seems to come later and then apparently as a separate vocabulary item independent of its physical meaning (i.e., something on the order of a homonym). The realization of the double-function property of these terms is the last thing to occur and then usually not spontaneously within the age groups studied.

There is, however, some problem to this type of interpretation given the procedure and stimuli used. The crux of the problem revolves around the specific double-function terms used. All of the terms used (at least all of the terms specifically mentioned in the text of their study) involve frozen metaphors which in fact do exist as separate lexical entries. So, for example, if we look in as old a dictionary as Webster's Universal Dictionary of the English Language (1937) under the entry "hard," we find that the fifth definition (out of fourteen) runs as follows: "unfeeling, not easily moved by pity,...severe, obdurate...as a hard landlord" (p. 766). In accordance with what we have said previously, dead metaphors (i.e., those that regularly appear in a dictionary) should be learned as separate lexical items and need not necessarily, in the minds of children, have any connection whatsoever with the meaning of the term as it might be applied to a physical referent. The question of when (at what age) and in what manner children come to use and understand living or non-frozen metaphors is, therefore, essentially left untouched by this study. Unfortunately, we cannot use Asch and Nerlove's trends with any certainty in describing the development of novel metaphors in children. Clearly a more adequate procedure needs to be found if we are to determine the growth of figurative language in children.

D. Purpose of the Present Program

Given the rather sparse and somewhat confusing literature dealing with the topic of metaphor and metaphoric usage, the specific purposes of the present study can be summarized as follows: (1) to determine the basal level of metaphorical usage in grades three, four, and five; (2) to evaluate a series of books designed to promote and augment figurative usage in these grades; (3) to determine the differences between groups exposed to this series and those using the normal curriculum of the school; and (4) to uncover factors in an individual child's test profile that might predispose him to differential metaphorical usage.

Although, as we have seen, much has been written on the need for instruction in figurative language in such areas as critical reading and thinking on all grade levels extending as far as the college years (Cooper, 1954; Newton, 1964; Upton, 1964a, 1964b), there has been surprisingly little research done on the development of metaphor and on methods of instruction. The only developmental study, that of Asch and Nerlove (1960), seems to have basic design flaws which render their conclusions equivocal. The present study is a first step leading to much needed research in this area.

Reports of successful teaching methods on the elementary school level have also been scarce. Hughes (1967) describes a brief simplified method for eliciting metaphors and similes from children. Although effective, this procedure would suffice for no more than one or two lessons, and would need to be incorporated into a larger sequential program of instruction.

Synectics, Incorporated has published a series of workbooks (Making It Strange, 1968) designed to increase figurative language usage and creative thinking in grades three through five. Although recent related work (Holstein, 1972a; 1972b) has shown that similar types of materials have been effective in classroom settings, there is little or no information published as to the effectiveness of these books. If these materials are successful, the question then becomes should we begin with grade three or can children younger than grade three also benefit from similar instruction?

II. Methods of Procedure

The present study, then, represents an attempt to determine developmental levels for figurative language in elementary school children and to evaluate the effectiveness of the Making It Strange series in increasing the use of figurative language by these children. In addition, we will make specific recommendations as to implications for future research.

A. Sample of Subjects

The subjects for this study were enrolled in West Hills Elementary School in Knoxville, Tennessee. This city school is in a middle to upper middle class white neighborhood. All the classes in the school were organized by the principal into heterogeneous groups on the basis of achievement and intelligence tests prior to the beginning of the academic year. Two classes per grade, for Grades three, four, and five, were chosen to be included in the sample. One class per grade was designated as the experimental group, while the other was the control group. Table I contains the means and ranges of scores for both intelligence and achievement tests for each grade and each class. As can be seen from Table I, the classes are comparable and well matched.

B. Test Materials

Since there are no available standardized tests designed to elicit figurative language per se, the test procedure used in this study had to be specially designed. Before arriving at a specific test procedure, some consideration had to be given to the nature of the test used and to its acceptance in a classroom setting. It was thought desirable to design a test that would be as naturalistic as possible, i.e., one that would not specifically ask for figurative usage; but would allow the children to use metaphor in a commonly occurring classroom situation. The test procedure actually used consisted of five composition topics, with accompanying motivating questions for each topic.[1]

The child was instructed to choose one topic and write a composition about it. The same test was used for the pretest and the posttest for all grade levels. On the posttest the child was told, by the teacher, to choose a topic that he had not written on previously, although this condition was not stringently enforced as primary concern was on the mode of expression rather than on the content expressed. It was felt that since the number of subjects was large, there would be sufficient sampling of all topics, across all grades, on both the pretest and the posttest to insure that whatever differences there were in individual topics would not affect group scores. A prototype copy of the composition test can be found in Appendix A.

C. Training Materials

The experimental training materials consisted of the Making It Strange books (one for each grade level). These soft-cover books contain exercises, games, and compositions designed to acquaint and familiarize the students with creative models of thinking and writing, and with figurative language. No specific figures of speech are introduced or defined; rather they are presented and practiced as

---

[1] Although the original plan called for the use of a number of different test situations, teacher resistance forced the use of only the single composition probe described in the text.

TABLE I

RANGE OF SCORES AND MEAN SCORES FOR INTELLIGENCE AND ACHIEVEMENT TESTS FOR ALL CLASSES

Grade		Intelligence Test Scores		Achievement Test Scores	
		Range	$\bar{X}$	Range Percentile	$\bar{X}$ Percentile
3	Control	96-135	115.24	11-99	65.76
	Experimental	102-135	117.69	24-99	73.00
4	Control	90-136	114.46	8-99	68.39
	Experimental	95-132	117.81	14-98	75.02
5	Control	94-135	116.73	29-97	69.33
	Experimental	99-134	116.75	9-99	70.30

games or exercises with which the children work as they proceed through the book.

The Making It Strange materials require very little "teaching," as such, by the teacher. Within the first few lessons on each grade level, the books stress that there are no right or wrong answers. The emphasis is upon what each child, as an individual, thinks about the particular lesson topic and why he thinks what he does. Sharing of one's own views and discussing the views of others are the major techniques employed.

D. Testing Procedures

After permission to conduct the study was granted by both the Board of Education and the Principal, the teachers of Grades three through five were contacted and asked to attend a meeting. At this meeting, the background and purposes of the study were carefully explained. The teachers then examined the test materials and the Making It Strange series. The timetable of the study and the responsibilities of the respective teachers were agreed upon.

The teachers were told to read the topics to the children after the children received their own copies of the test. The children were then told to choose one topic and to write a composition of at least one page on that topic. The tests were all given on the same day. The Making It Strange books were used by the experimental teachers and classes every day for a period of four weeks (twenty instructional days). The posttest was given in the same manner as the pretest on the specified date.

All teachers were asked not to teach figurative language before the pretest and during the time period of the experiment. They were strongly assured that they themselves as teachers were not being evaluated.

Each child's score on the posttest was compared with his own score on the pretest. The control groups were included to assess any changes that might occur with the simple passage of time.

E. Rater Training and Evaluation Procedures

Since the pretest and posttest were experimenter designed as a means of eliciting metaphorical usage, no standardized grading system was available. It became necessary to train raters to identify figurative usage. A manual developed by Barlow, Kerlin, and Pollio (1971) for the purpose of identifying figures of speech in psychotherapy tapes was used for this purpose. Three raters were trained to recognize all the standard rhetorical figures of speech indicated in the manual. After completion of training, each rater independently rated each class and wrote down the figures of speech for each child in each class, one class at a time. Raters' judgments were tallied on the basis of the following coding scheme:

3 + 0: This means that all three raters independently judged this instance as figurative;

2 + 1: this means that two of the three raters independently judged the instance to be figurative and after discussion the third rater agreed;

1 + 2: this means that only one of the raters independently judged the instance to be figurative and after discussion the other two raters agreed;

2 - 1: this is the case in which two raters independently chose an instance as figurative but the third judge after discussion did not agree;

1 - 2: this is the case where one rater independently chose an instance as figurative but the other two raters still disagreed after discussion.

Each rater also noted whether the figure of speech was either frozen (F) or novel (N). Thus by using this system, an instance might be coded, 1 + 2F. This would be a case in which the instance was first independently chosen by only one rater as figurative and after discussion the other raters agreed. This code also reveals that the instance was judged to be frozen. In accordance with a suggestion contained in the manual, no rater worked for more than one-half to three-quarters of an hour at each sitting.

The coding scheme outlined above was used to determine the pattern of agreement and disagreement between the raters. Table II presents these data for all three grades. Probably the best way in which to read this table is from the bottom up. For Grade 3 all raters scored a total of 168 units; for Grade 4, they scored 205 units; and for grade 5, 253 units. Of the 168 units scored for Grade 3, 150 (88%) were agreed upon by all three raters after their discussion sessions, while 18 (12%) were never agreed upon. For Grade 4, raters agreed 85 percent of the time, while for Grade 5, raters agreed 82 percent of the time. Of the 150 units agreed upon for Grade 3, there were a total of 83 frozen figures and 67 novel ones; for Grade 4, there were 111 frozen and 67 novel; while for Grade 5, there were 142 and 72 respectively.

An examination of these judgments shows that for all three grades, raters found it easier to pick out novel than frozen figures. For Grade 3, the proportion of 3 + 0 and 2 + 1 judgments was 0.73 for novel figures and 0.54 for frozen. The comparable values were 0.70 and 0.58 for Grade 4; and 0.74 and 0.53 for Grade 5. What this means is that raters tended to miss frozen metaphors more frequently than novel ones, with about 45 percent of the frozen metaphors being noticed by only a single rater during his rating of the compositions. Although agreement values for the 2 + 1 and 3 + 0 condition reached about a 75 percent level for novel and a 53 percent level for frozen figures, these "agreement values" are clearly very conservatively figured. Raters also agreed very frequently on the nonoccurrence of metaphors, so that even the higher values of 88%, 85%, and 82% probably represent underestimates of how well raters actually agreed in doing this task.

A similar agreement table was computed for the posttest compositions. In general the results shown in Table III parallel in all important respects those presented in Table II. The percentage agreement scores were slightly higher for the posttest than for the pretest in Grades 4 and 5 and slightly lower in Grade 3.

In order to determine if there were any significant differences in mean number of figures of speech detected by each rater, a random sample of five children was drawn from the total grade rated that week and an analysis of variance computed over these values. Table IV shows the results of the analysis of variance for Grade 3. Although only the between rater variance was of interest, the between student variance was calculated in order to account for as much of the total variance as possible, and to provide an appropriate error term. As can be seen from Table IV, the between rater scores did not differ significantly. As might be expected, individual differences between children produced a significant F-ratio at better than the 0.05 level.

Similar analyses were computed for between rater judgments and individual differences for five random subjects for Grade 4 and Grade 5. Table V contains the analysis for Grade 4. These results show a large individual difference factor ($F = 24.91$, $p<0.01$); but no significant rater differences.

TABLE II

PATTERN OF RATER AGREEMENTS FOR SCORING AND FIGURATIVE LANGUAGE CATEGORIES OVER ALL THREE GRADES - PRETEST

Rater Scoring Category	Grade											
	Third				Fourth				Fifth			
	Frozen		Novel		Frozen		Novel		Frozen		Novel	
	N	%	N	%	N	%	N	%	N	%	N	%
3 + 0	22	26	31	46	30	27	18	27	31	22	31	43
2 + 1	23	28	18	27	34	31	29	43	44	31	22	31
1 + 2	38	46	18	27	47	42	20	30	67	47	19	26
Subtotal	83		67		111		67		142		72	
Subtotal Accepted	150		88%		178		85%		214		82%	
1 - 2	18				27				38			
2 - 1	0				0				1			
Subtotal Rejected	18		12%		27		15%		39		18%	
Total Units Scored	168		100%		205		100%		253		100%	

TABLE III

PATTERN OF RATER AGREEMENTS FOR SCORING AND FIGURATIVE LANGUAGE CATEGORIES OVER ALL THREE GRADES - POSTTEST

Rater Scoring Category	Grade											
	Third				Fourth				Fifth			
	Frozen		Novel		Frozen		Novel		Frozen		Novel	
	N	%	N	%	N	%	N	%	N	%	N	%
3 + 0	32	31	21	33	28	23	53	52	21	23	33	45
2 + 1	36	35	26	41	52	42	24	24	32	34	26	35
1 + 2	34	33	17	26	44	35	24	24	40	43	15	20
Subtotals	102		64		124		101		93		74	
Subtotal Accepted	166		84%		225		89%		167		90%	
1 - 2	31				22				19			
2 - 1	1				6				0			
Subtotal Rejected	32		16%		28		11%		19		10%	
Total Units Scored	198		100%		253		100%		186		100%	

TABLE IV

ANALYSIS OF VARIANCE BETWEEN RATER JUDGMENTS
OF FIVE RANDOM STUDENTS, GRADE THREE

Source	SSD	df	MS	F
Between Raters	0.53	2	0.265	0.996
Between Students	6.66	4	1.665	6.10*
Residual Error	2.13	8	0.266	
Total	9.33	14		

*p < 0.05

TABLE V

ANALYSIS OF VARIANCE BETWEEN RATER JUDGMENTS
OF FIVE RANDOM STUDENTS, GRADE FOUR

Source	SSD	df	MS	F
Between Raters	0.133	2	0.067	0.288
Between Students	22.933	4	5.730	24.910**
Residual Error	1.867	8	0.233	
Total	24.933	14		

**p < 0.01

TABLE VI

ANALYSIS OF VARIANCE BETWEEN RATER JUDGMENTS
OF FIVE RANDOM STUDENTS, GRADE FIVE

Source	SSD	df	MS	F
Between Raters	1.60	2	0.80	0.51
Between Students	16.40	4	4.10	2.64
Residual Error	12.40	8	1.55	
Total	30.40	14		

Table VI presents the comparable analysis for Grade 5. It can be noted that in Grade 5 neither the between rater nor the between student F-ratios were significant. These results also showed that the between rater judgment for Grade 4 produced an F-ratio of 0.29, while the between rater F-ratio for Grade 5 was 0.51. In neither case were the between rater F-ratios significant.

The conclusion reached from the data presented in Tables IV through VI was that there were no significant differences among the total figures rated by each of the three raters for these randomly selected protocols, and that raters did not differ in their ability to detect figurative language.

III. Analysis of the Data

In order to assess the place of figurative language in the compositions of elementary school children, many statistical analyses were conducted on the rated data of both the pretest and the posttest.

A. Instances of Figurative Usage: Pretest

For each child, the number of figures per composition was converted to a percentage score, i.e., the number of metaphors divided by the total number of words in the composition multiplied by 100. This value, in effect, represents the number of metaphors per hundred words of text. Following this conversion, means were computed for all three grades for frozen and novel metaphors separately.

These results are presented in Figure 1 where it can be seen that students produced a larger number of frozen than novel metaphors in each of the three grade levels. Means for Grade 3 were Frozen = 1.69, Novel = 1.31. Grade 4 means were Frozen = 1.20, Novel = 0.82; and Grade 5 means were Frozen = 1.44, Novel = 0.72. As Figure 1 also shows, the number of metaphors decreased over successive grade levels, although this decrease was more marked for novel than for frozen figures.

Given these trends, an analysis of variance was computed over the scores. Since there was an unequal number of students in each of the classes, some corrections had to be made before this statistical analysis could be carried out. In order to correct for an unequal number of students in each class, the class with the smallest number of cases (N = 23) was chosen as the standard, and cases were dropped randomly from all other classes until there were only 23 students in the remaining five classes. In dropping cases, care was taken to equate the proportion of zero scores remaining with the proportion found in the original sample. So, for example, if a class had 35 scores and 7 were zero, 20 percent of the final 23 scores selected also have a value of zero (i.e., between four and five).

Once this correction had been applied, the results of an analysis of variance computed over these data (see Table VII) showed that students produced a significantly larger number of frozen than novel figures: $F_{1,132} = 7.76$; $p<0.01$, and that the total percentages of figures dropped significantly over the three grades considered: $F_{2,132} = 3.80$; $p<0.05$.

B. Intercorrelations Between Pairs of Response Measures

Part of the reason for expressing the number of figures produced as a percentage was to correct for unequal composition length. But individual students within each grade also produced compositions of unequal length, and it seemed reasonable to ask if there was any relationship between composition length and the number of novel figures produced, and composition length and the number of frozen figures produced. In addition, it also seemed reasonable to determine if students who

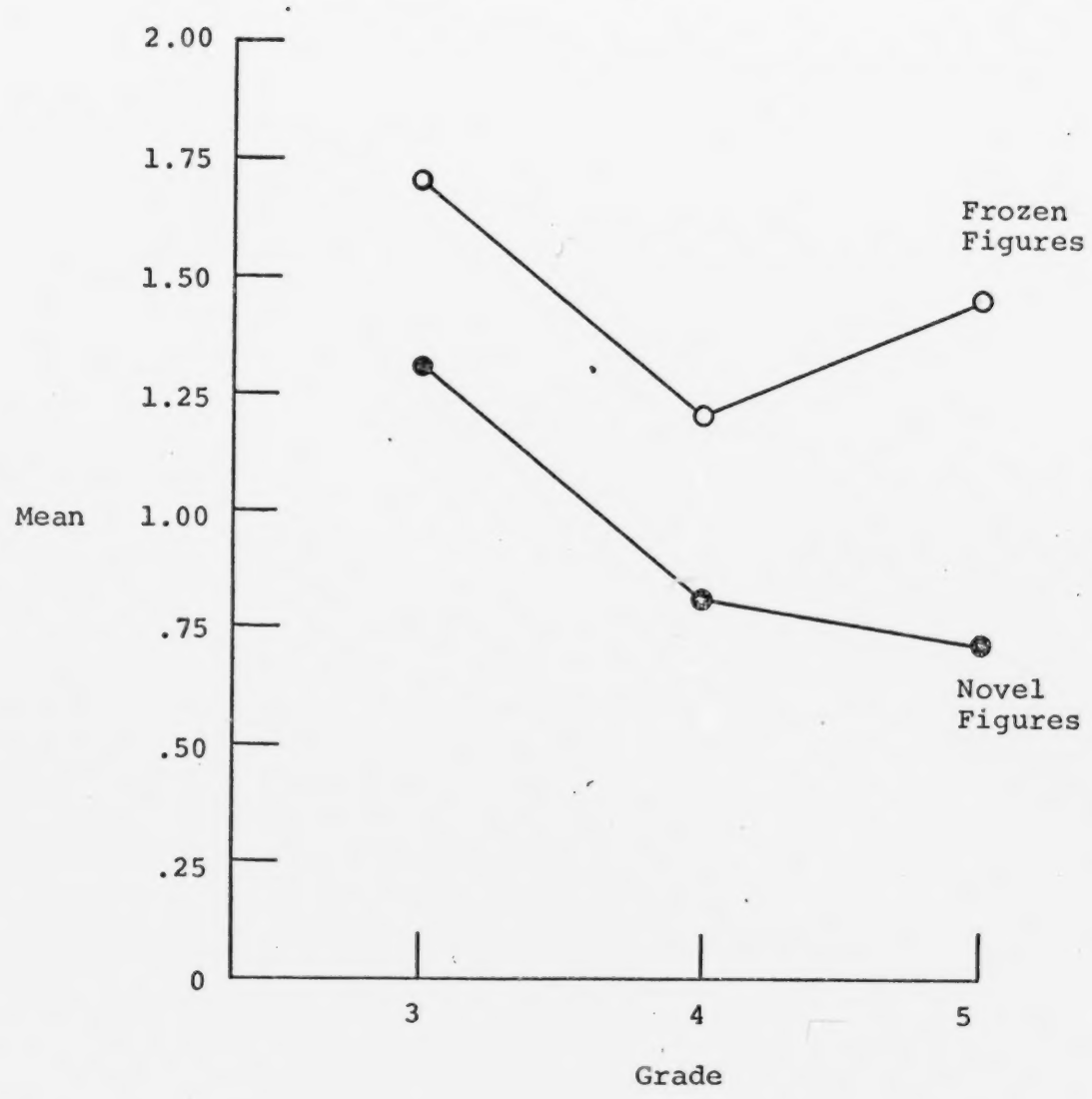

FIGURE 1

MEAN PERCENTAGE SCORES FOR NOVEL AND FROZEN FIGURES FOR ALL THREE GRADE LEVELS - PRETEST

TABLE VII

ANALYSIS OF VARIANCE OF PERCENTAGE SCORES - PRETEST

Source	SSD	df	MS	F
Between Subjects	248.02	137		
Grade Level (B)	12.85	2	6.43	3.80*
Control vs. Experimental (C)	3.81	1	3.81	2.25
B x C	8.63	2	4.32	2.56
Error	222.73	132	1.69	
Within Subjects	305.33	138		
Frozen vs. Novel (A)	16.81	1	16.81	7.76**
A x B	1.82	2	0.91	0.42
A x C	0.38	1	0.38	0.18
A x B x C	0.22	2	0.11	0.05
Error	286.10	132	2.17	
Total	553.35	275		

*$p < 0.05$

**$p < 0.01$

produced a large number of novel figures also produced a large number of frozen figures. Table VIII presents the pattern of intercorrelations for all three grades as well as the correlations for all grades combined.

Looking first at the combined correlations, it can be seen that total words and frozen figures correlated more strongly than total words and novel figures (0.49 to 0.27). The correlation of 0.19, between frozen and novel figures, while marginally significant, is really too small to be of great theoretical importance. A further examination of such correlations for all three classes essentially confirms this conclusion: in two of three classes, the correlation between frozen and novel figures was almost zero. An examination of the other two types of correlation produced results similar to those reported for the combined grades; significant and high positive correlations between words and frozen figures, and smaller correlations between words and novel figures. There is some small tendency for these latter two types of correlations to increase over successive grades.

C. Instances of Figurative Usage: Posttest

Percentage scores were computed in the same manner for the posttest figures of speech as for the pretest figures. Means for the posttest percentage scores for novel and frozen figures are shown in Figure 2. The means for the three grades were: Grade 3, Frozen = 2.53, Novel = 1.60; Grade 4, Frozen = 1.49, Novel = 1.66; Grade 5, Frozen = 1.58, Novel = 1.02. The means for each category, for each grade, increased over their pretest values.

Figures 1 and 2 can be further combined and partitioned to allow for a comparison of the pretest versus the posttest values for both control and experimental groups for the frozen and novel categories of figurative language. Figure 3 presents this comparison for novel usage. The means presented in this figure were obtained after all necessary cases were dropped because of either experimental mortality, i.e., subjects who either moved or were excessively absent and did not take the posttest, or because of the need to equalize the number of Ss in each group. In this latter case, subjects were dropped randomly.

Figure 3 shows that while the control groups, which on the pretest had a higher percentage of novel usage (Grade 3 = 1.60, Grade 4 = 0.80, Grade 5 = 0.[illegible]), remained approximately the same (Grade 3 - 1.50, Grade 4,= 1.46, Grade 5 = 0.83); the experimental groups improved markedly in all three grades (pretest: Grade 3 = 0.77, Grade 4 = 0.61, Grade 5 = 0.54; posttest: Grade 3 = 1.71, Grade 4 = 1.86, Grade 5 = 1.21).

Figure 4 presents results for the same groups for frozen usage. The means used in this figure were obtained in the same manner as they were for novel usage, i.e., for Figure 3; and students involved in both comparisons were the same. As can be seen from this figure, subjects in both experimental and control groups increased in their use of frozen figurative language from the pretest (Grade 3: Control = 2.15, Experimental = 1.37; Grade 4: Control = 1.08, Experimental = 1.43; Grade 5: Control = 1.30, Experimental = 1.26) to the posttest (Grade 3: Control = 3.53, Experimental - 1.52; Grade 4: Control = 1.17, Experimental = 1.81; Grade 5: Control = 1.43, Experimental = 1.74). The only group whose results differed markedly was the Grade 3 control group, which showed both an unusually high pretest and an unusually high posttest.

In order to analyze these results further, analyses of variance were computed for pretest versus posttest scores on novel figurative usage and on frozen figurative usage separately. Table IX contains the results of the analysis for novel usage. The results presented in this table show that there was a significant increase in novel figurative usage from the pretest to the posttest ($F_{1,108} = 10.65$, $p<0.01$) and that there was a significant interaction between the pretest and the posttest and the experimental and control groups ($F_{1,108} = 5.36$, $p<0.05$). Figure 3 presents this interaction graphically. Although all groups did increase

TABLE VIII

PATTERN OF INTERCORRELATIONS BETWEEN PAIRS OF MEASURES FOR ALL THREE GRADES SEPARATELY AND TOGETHER - PRETEST

Correlation Between:	Grade 3	Grade 4	Grade 5	Combined
Number of Words and Frozen Figures	0.42**	0.39**	0.62**	0.49**
Number of Words and Novel Figures	0.29*	0.23*	0.40**	0.27**
Number of Frozen and Novel Figures	0.04	0.34**	0.14	0.19*
Number of Cases	53	62	59	174

*$p < 0.05$

**$p < 0.01$

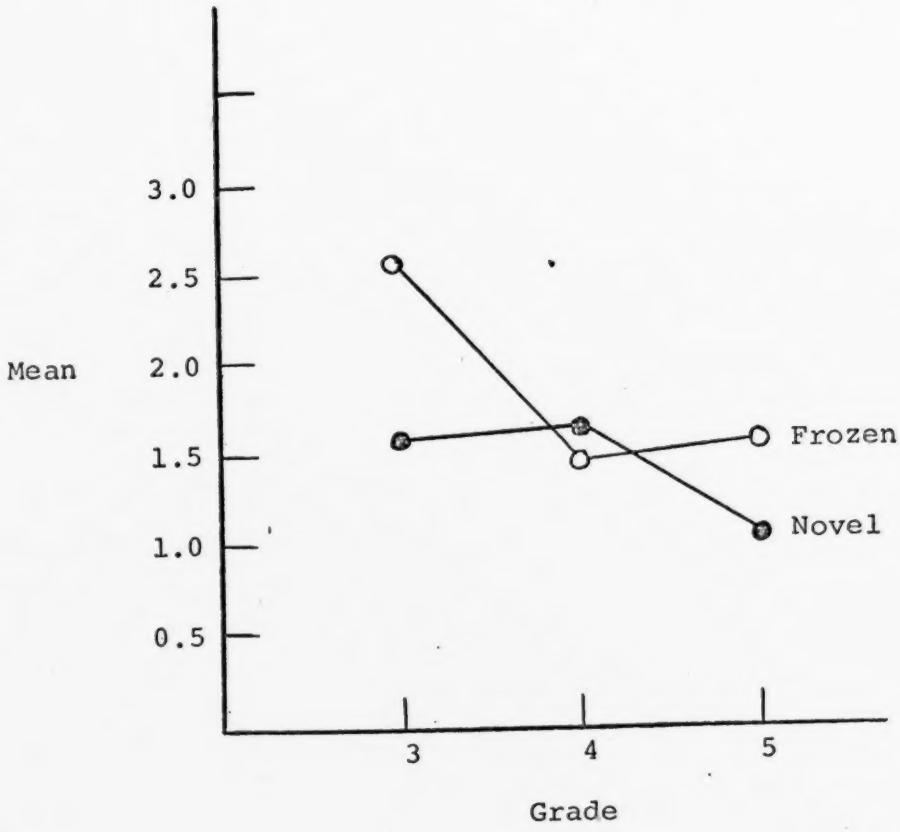

FIGURE 2

MEAN PERCENTAGE SCORES FOR NOVEL AND FROZEN FIGURES
FOR ALL THREE GRADE LEVELS - POSTTEST

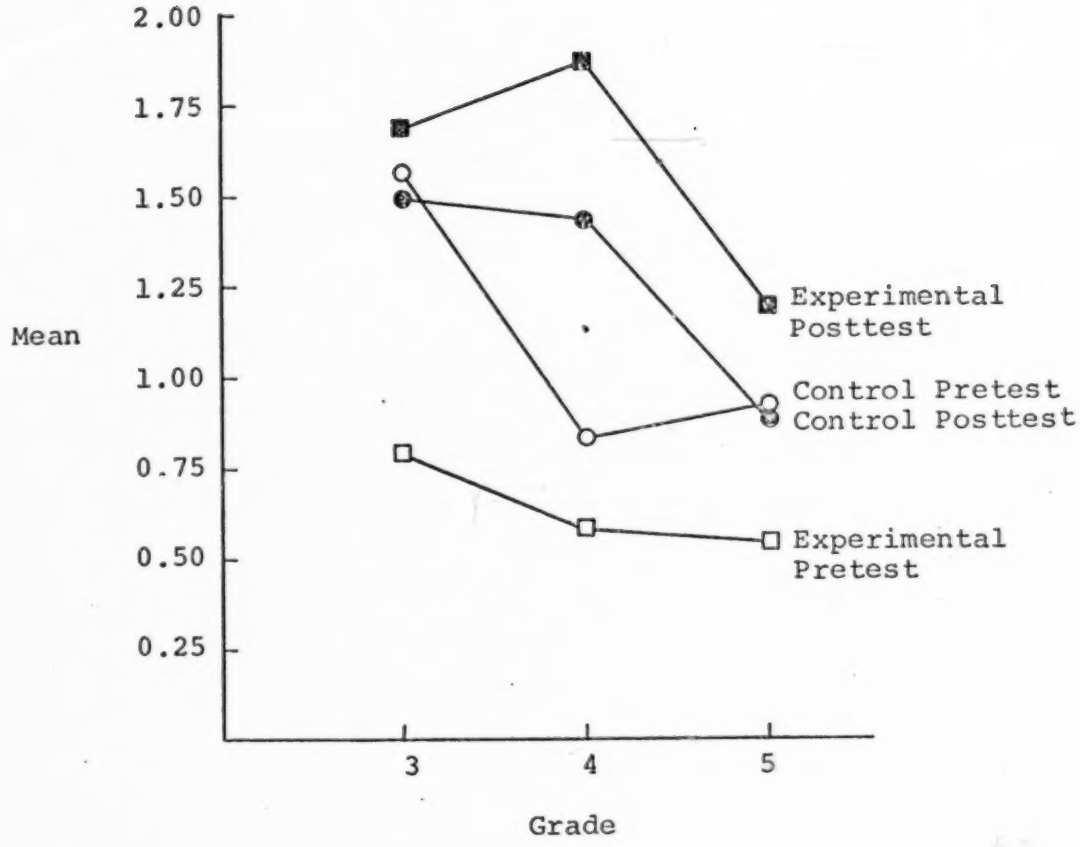

FIGURE 3

PRETEST VERSUS POSTTEST OVER ALL GRADES, CONTROL VERSUS EXPERIMENTAL - NOVEL USAGE

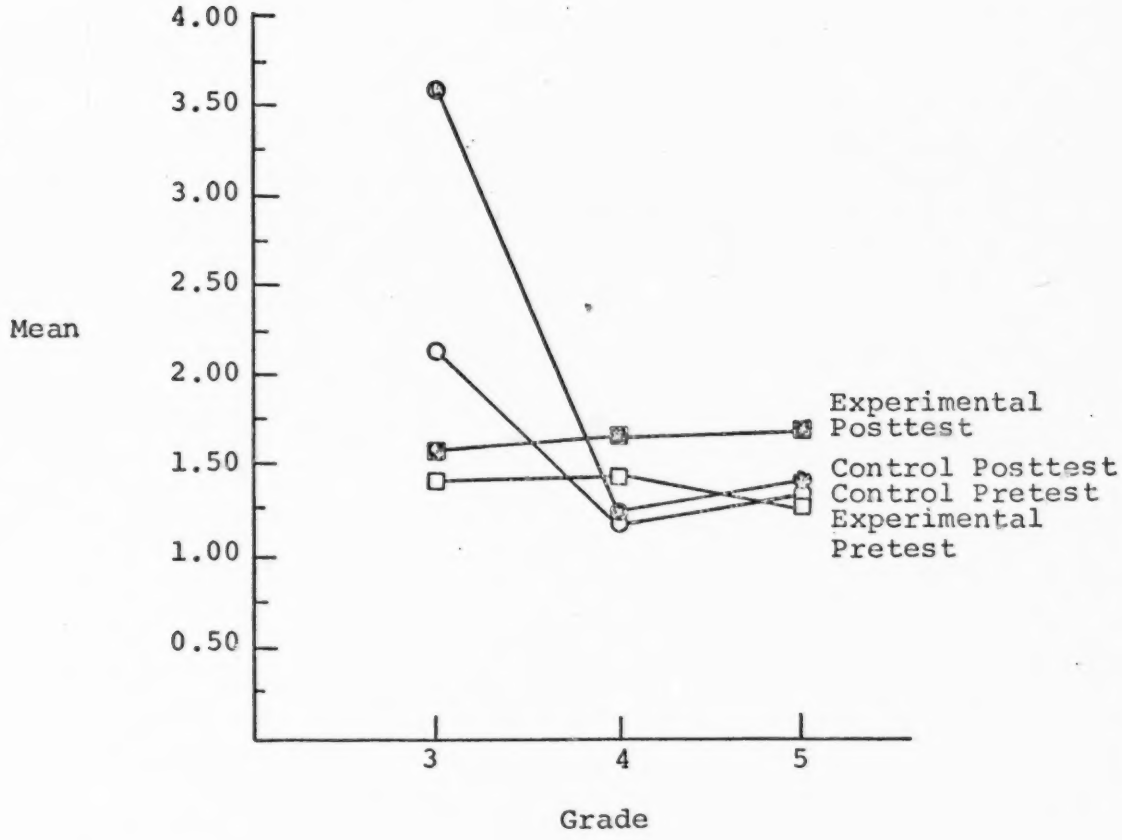

FIGURE 4

PRETEST VERSUS POSTTEST OVER ALL GRADES, CONTROL VERSUS EXPERIMENTAL - FROZEN USAGE

TABLE IX

ANALYSIS OF VARIANCE PRETEST VERSUS POSTTEST - NOVEL USAGE

Source	SSD	df	MS	F
Between Subjects	312.02	113		
Grade Level (B)	10.52	2	5.26	1.90
Control vs. Experimental (C)	0.23	1	0.23	0.08
B x C	1.82	2	0.91	0.33
Error	299.45	108	2.77	
Within Subjects	212.66	114		
Pretest vs. Posttest (A)	17.79	1	17.79	10.65**
A x B	4.58	2	2.29	1.37
A x C	8.95	1	8.95	5.36*
A x B x C	0.49	2	0.25	0.15
Error	180.85	108	1.67	
Total	524.68	227		

*$p < 0.05$

**$p < 0.01$

in novel usage, the experimental groups showed a markedly larger gain. The pretest level for the control groups was initially higher than for the experimental groups (although not significantly so: $F_{1,108} = 2.25$; see Table VII) while the posttest level for the experimental groups was higher than that of the control groups.

A comparable analysis of variance for frozen usage is shown in Table X. As can be seen from the results presented in this table, there is a significant decrease in usage over grades ($F_{2,108} = 5.99$, $p<0.01$) and a significant interaction between grade level and experimental treatment groups ($F_{2,108} = 8.18$, $p<0.01$). The reason for this statistical result can be seen in Figure 4: all groups increased in frozen usage from pretest to posttest, although Grade 3 control group, which was initially quite high, showed a much greater increase than any of the other groups. The scores obtained by students in this group would seem to account for the significant effects reported in Table X.

A further analysis using a Matched Pairs Sign Test was done in order to see if there was a significant directional change in usage from the pretest to the posttest. In order to use this test, difference scores were computed for each subject used in the previous analysis. Since the posttest scores were higher, the difference score was computed by subtracting the pretest score from the posttest score. A positive number indicated an increase in figurative usage, while a negative number indicated a decrease in figurative usage. The results of the Sign Test are presented in Table XI. As these results show, there was a significant number of students who increased their use of novel figures of speech in both the fourth and fifth grades experimental groups and in the total experimental group novel category, collapsed over grades. No frozen figure category was significant. These difference scores, when presented graphically, illustrate this point, and help clarify Figures 3 and 4. Figure 5, therefore, presents the mean difference scores for novel usage. As can be seen, the experimental groups for all three grades showed a greater increase in novel usage on the posttest (difference score means, Grade 3 = 0.94, Grade 4 = 1.25, Grade 5 = 0.67) than did the control groups (difference score means, Grade 3 = -0.10, Grade 4 = 0.65, Grade 5 = -0.07).

Figure 6 presents the mean difference scores for frozen usage. In contrast to results for novel figurative usage, results for frozen usage show very little difference between experimental and control groups (difference score means, experimental groups, Grade 3 = 0.15, Grade 4 = 0.38, Grade 5 = 0.47; control groups, Grade 3 = 1.38, Grade 4 = 0.09, Grade 5 = 0.15).

D. Intercorrelations Between Pairs of Response Measures

In order to determine if the pattern of intercorrelations between pairs of response measures remained the same on the posttest as on the pretest, similar correlations to those computed on the pretest data were computed on the posttest data. Table XII presents these intercorrelations. Looking first at the combined correlations, results show that number of words and frozen figures correlated more strongly than did number of words and novel figures (0.37 to 0.21). This is the same pattern that occurred on the pretest, although the correlations are of smaller magnitude on the posttest. The correlation between number of frozen and number of novel figures was not significant on the posttest in the combined grades. This correlation was only marginally significant in the fifth grade and not significant in the other two grades. The correlations for each grade are, however, more equivocal on the posttest than on the pretest; although the total pattern of correlations remains somewhat the same.

TABLE X

ANALYSIS OF VARIANCE PRETEST VERSUS
POSTTEST - FROZEN USAGE

Source	SSD	df	MS	F
Between Subjects	321.435	113		
Grade Level (B)	27.91	2	13.955	5.99**
Control vs. Experimental (C)	3.63	1	3.63	1.56
B x C	38.13	2	19.065	8.18**
Error	251.765	108	2.33	
Within Subjects	320.685	114		
Pretest vs. Posttest (A)	10.78	1	10.78	3.91
A x B	3.22	2	1.61	0.58
A x C	0.58	1	0.58	0.21
A x B x C	7.59	2	3.795	1.38
Error	298.515	108	2.76	
Total	642.12	227		

**$p < 0.01$

TABLE XI

SIGN TEST OVER ALL THREE GRADES, CONTROL VERSUS EXPERIMENTAL - FROZEN AND NOVEL

Condition		Three		Four		Five		Total	
		Frozen	Novel	Frozen	Novel	Frozen	Novel	Frozen	Novel
Experimental	number of positive instances[a]	9	11	10	14**	10	13*	29	38*
	N[b]	17	16	17	17	17	18	51	51
Control	number of positive instances[a]	13	5	10	9	10	7	33	21
	N[b]	19	14	17	15	19	17	55	46

*$p < 0.05$

**$p < 0.01$

[a]A positive instance is one in which the posttest score is larger than the pretest score.

[b]N values are uneven due to the dropping of all zero difference cases as required by a sign test.

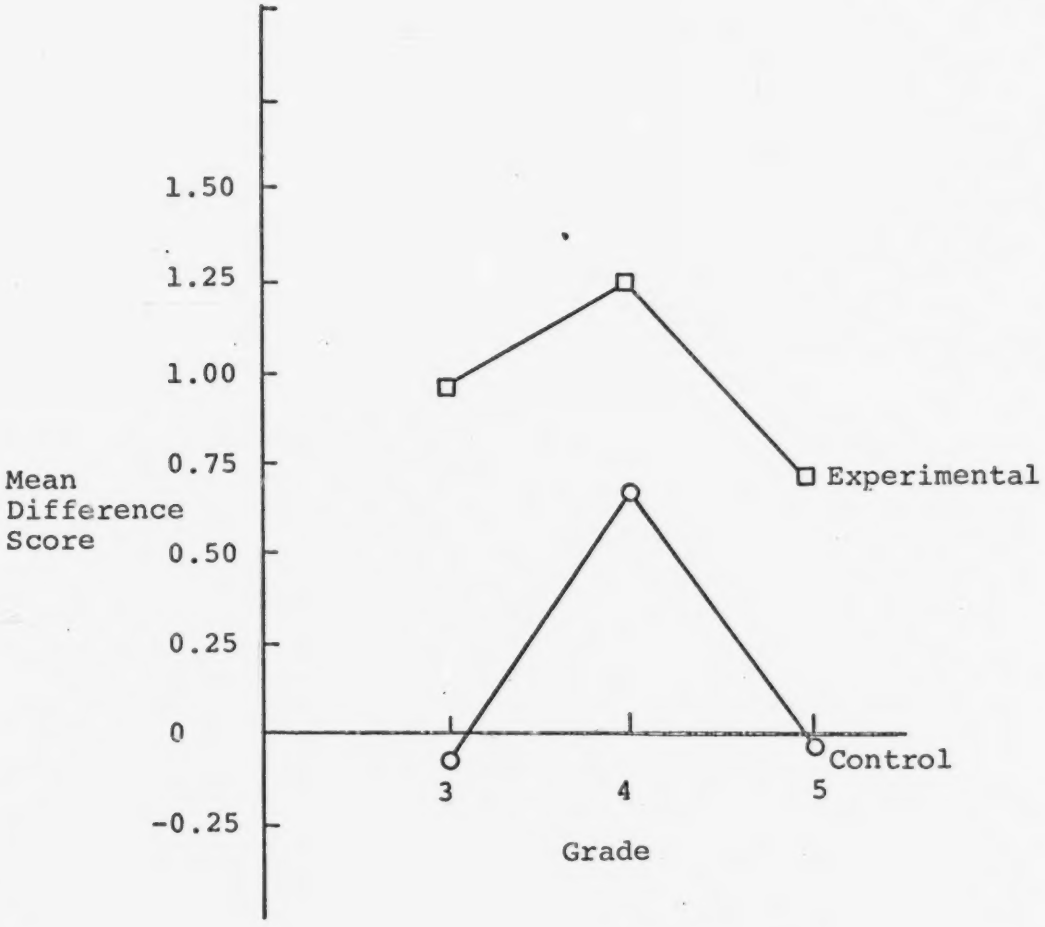

FIGURE 5

MEAN DIFFERENCE SCORES FOR NOVEL FIGURATIVE USAGE OVER ALL THREE GRADES - CONTROL VERSUS EXPERIMENTAL

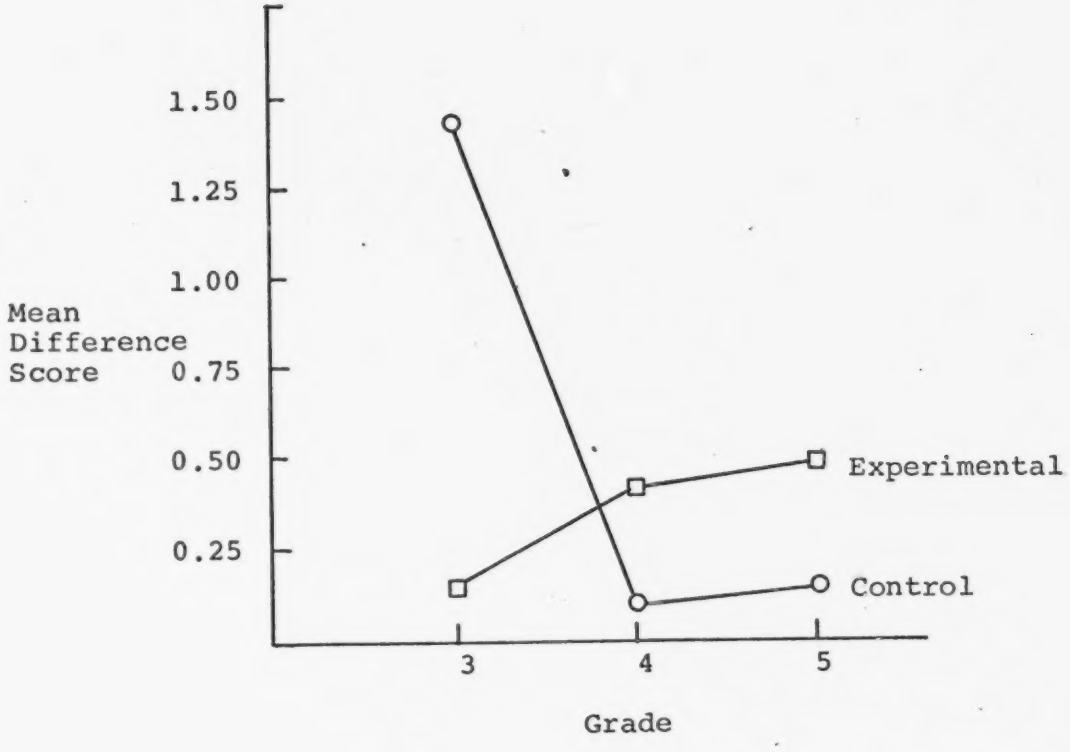

FIGURE 6

MEAN DIFFERENCE SCORES FOR FROZEN FIGURATIVE USAGE OVER ALL THREE GRADES - CONTROL VERSUS EXPERIMENTAL

TABLE XII

PATTERN OF INTERCORRELATIONS BETWEEN PAIRS OF MEASURES FOR ALL THREE GRADES SEPARATELY AND TOGETHER - POSTTEST

Correlation Between:	Grade 3	Grade 4	Grade 5	Combined
Number of Words and Frozen Figures	0.27*	0.57**	0.20	0.37**
Number of Words and Novel Figures	0.33*	0.14	0.20	0.21*
Number of Frozen and Novel Figures	0.14	0.06	0.27*	0.13
Number of Cases	38	38	38	114

*$p < 0.05$

**$p < 0.01$

IV. Discussion

A. Pretest

Contrary to conclusions suggested by Asch and Nerlove (1960), children in the present study were able to produce a substantial number of metaphors -- both novel and frozen -- even as early as third grade. As a matter of fact, third grade children seemed to produce a greater amount of figurative language than children in the higher grades.

How are these differences between the results presented here and Asch and Nerlove's trends to be understood? Probably the most important difference concerns the task used to measure a child's understanding of figurative language. Asch and Nerlove were interested in seeing whether a child could explain why terms such as hot or sweet could refer to both physical and psychological aspects of things and people. In contrast, the interest in this study was in seeing if children could make use of figurative language rather than in seeing if they could describe the "whys" and "wherefores" of such language. It is not an uncommon finding that the use of a language skill often surpasses a speaker's ability to describe what's going on. In the case of grammar, for example, we can all sense the difference between (They) (are flying) (planes), and (They) (are) (flying planes); yet very few people could draw or even describe the appropriate phrase structure. As noted earlier, part of the reason for the child's inability to explicate the metaphoric relationships inherent in double-function terms may be because such terms are frozen into the language. Under this condition it is not at all surprising that young children should consider such terms more as homonyms than as semantically related words.

Using children's compositions as a source of data may also help account for the decrease in the number of metaphors produced over successive grades. Informal examination of these compositions showed very little change in vocabulary used, but similar to results found by Loban (1963), profound changes in the child's control of grammar and spelling. The composition task seems to be one in which a child is very strongly concerned about getting a good grade, and that means: don't rock the boat. Experimentally this may mean that a more accurate way in which to assess developmental trends in connected discourse would be to have children speak their "compositions" rather than to write them. In this way, it might be possible to remove the demand characteristics that go along with writing a composition in an elementary school classroom. In any event it is interesting to note that in the present context the composition task is one in which the child uses progressively less figurative language as he gains more experience in writing.

Another factor that must be considered in interpreting these results is the teacher. Although all teachers were assured that they were not being evaluated, and were asked not to teach figurative language as such either prior to or during the experimental period (unless it occurred within the normal curriculum such as in the reader or English text), we have no way of knowing whether these directions were followed. Some teachers may have primed their classes prior to either or both the pre and posttests. Since such priming would probably be confined to examples and explanations of figurative usage, it would include primarily frozen figures. If so, the results of the analyses on frozen figurative language would be more affected than the analyses of novel usage. It would be difficult to provide practice on and to prime the use of novel metaphors with a few simple examples. This teacher priming may account for the results found in the analyses of frozen usage, although there is no way in which to secure unequivocal proof. A larger sample with more classes on each grade level for each condition could serve to distribute such effects randomly across all conditions.

Given these limiting conditions, what conclusions can be drawn in regard to figurative language? Here let us turn to the correlational data where the major finding was a strong correlation between number of words and frozen figures. Although it is difficult to interpret correlations unequivocally, these data do seem to imply that frozen metaphors function pretty much as regular vocabulary items and that the greater the written output, the greater the number of frozen figures. One other piece of data also suggesting that frozen figures should be considered largely as lexical items can be found in the relatively greater proportion of 1 + 2 ratings for frozen as opposed to novel figures. What this means is that all but one of the raters often missed such usage, but that once it was pointed out, both other raters readily agreed. Such usage, in contrast to novel figures, is much less compelling and consequently much more easily missed.

The essentially zero correlations between number of frozen and novel figures simply means that subjects who produce a large number of novel metaphors do not necessarily produce a large number of frozen metaphors. If the appearance of novel metaphors in compositions can be considered as some indication of both metaphoric thinking and lack of concern over grades, and if frozen metaphors represent simple lexical choices, then no correlation should be expected. Using frozen figures is probably irrelevant to determining whether or not a child can think metaphorically and, by extension, creatively.

What about the small, but still significant, correlations between novel usage and composition length? Here again the issue revolves around the meaning of correlation; and we may speculate that one property of novel usage is to promote greater interest and thereby increase the length of the composition produced. Unlike frozen figures, novel figures may serve to bring about longer compositions and are not simply a by-product of longer compositions. That is, of course, a post hoc explanation; and before we would put any confidence in it, we would have to test it further, i.e., by seeing if children presented with highly metaphoric composition fragments would produce longer completions than students presented with composition fragments devoid of figurative language. In any case, the experiment can be done, and such results should help to clarify the relationship of novel metaphors to composition length.

B. Posttest

In keeping with the results of the pretest on number of metaphors used, similar results were found on the posttest, i.e., grade school children do seem to produce figurative language in compositions and this production seems to decrease as school grade increases.

The questions raised for the posttest, however, must necessarily be different than those raised for the pretest. The posttest was designed to evaluate the effectiveness of a specific training procedure, the Making It Strange series, in augmenting the production of metaphor. The results seem to indicate that this series has no effect on the use or production of frozen figurative language. The results concerning frozen usage make sense if again we consider frozen figures simply as lexical items. A training procedure designed to initiate "strange" or unusual patterns of viewing the world or thinking about it should have little or no effect on vocabulary choice. As present results indicate, the Making It Strange series did not influence this use. The sharp decrease between third grade and the other two grades in the use of frozen figures can, perhaps, again be attributed to the composition task itself. A frozen figure may be considered by the child as a colloquial or even slang-like expression and the "get a good grade, don't rock the boat" philosophy may prevail.

In contrast to these results for frozen usage, the Making It Strange series did affect the production of novel figures of speech. This fact, however, seems to be slightly influenced by grade level with the fourth and fifth grades showing

a significant increase, and the third grade only a marginally significant increase. Given these results, what can be said about this series of workbooks? Perhaps the best way in which to answer this question is in terms of two further questions: Does this series, in training a child to think differently, make him more aware of unusual relationships or does it give him tacit permission to exhibit this "strange" use of language in a restrictive composition task? Does the use of this series teach a process of looking at the world metaphorically or does it free school-age children from the "don't rock the boat" philosophy? If the former is true, then there is no question of the value of the series in promoting metaphoric thinking. If the latter is true, then would one or two short lessons by the teacher, or, more simply, the reinforcement of unusual language behavior by the teacher produce the same results? On the basis of the present study it is impossible to specify exactly how the Making It Strange books brought about these results. A further series of studies would be necessary to answer these questions, where it would be possible to pit direct reinforcement procedures against the Making It Strange books themselves.

Turning now to the correlational data, it can again be noted that frozen figurative usage correlates significantly with length of composition, as does novel figurative usage and length of composition, at least across all grades; the same conclusions drawn about these correlations on the pretest can be drawn about these correlations on the posttest; viz., frozen figures are an accidental concomitant of composition length, while novel usage may in fact bring about longer compositions. Only future research could help explicate these correlations further.

V. Conclusions and Recommendations for Future Research

Figurative language has often been considered as an ornament in creative writing. In recent years the importance of this device as a heuristic has been stressed. Unfortunately, there are little developmental data in this area. The teaching of figurative language in the elementary schools has also been largely neglected. The present study attempted to determine the frequency of figurative language in children's compositions and to evaluate a program designed to increase the occurrence of figurative language in children in the elementary schools.

A. Conclusions

On the basis of the results of the present study the following conclusions seem reasonable:

1. Children use figures of speech (both novel and frozen) in their compositions as early as the third grade, and novel usage seems to decrease over grade level.

2. The experimental program (Making It Strange) increased the frequency of occurrence of novel figurative usage in the children exposed to this series as compared to their initial output, and as compared to the control groups. Exposure to this series, however, had no effect on the production of frozen figures of speech.

3. There is a strong relationship between frozen figures and length of composition, thereby suggesting that frozen figures might be considered simply as learned vocabulary items.

B. Implications for Future Research

Many questions can be raised as to the meaning of the results produced by this study. These questions suggest the following directions for future research:

1. A study similar to the present one should be conducted using a larger, more diversified sample of children, e.g., children from different sociocultural, socioeconomic, and geographic backgrounds. More teachers in each category (control and experimental), selected randomly, for each grade and a more extended time period would also be recommended. A study such as this might clarify developmental levels and enable broader generalizations to be made. The use of several teachers for each condition would control for varying teaching ability and for "priming" of the children in metaphoric processes prior to and during the experimental period.

2. A study similar to the one suggested above might be attempted with alterations of the pretest and posttest materials. The first alteration suggested would be the substitution of oral story telling for composition writing. This might help explicate the nature of the composition task itself, while still preserving the naturalistic quality of the probe. It might also free the children with the "don't rock the boat" philosophy of grades to use more unusual or creative language.

Further alterations of the pretest and posttest materials might include tasks which would attempt to tap metaphoric processes directly, e.g., analogy tasks, unusual uses of words, etc. These could be presented in either written or oral form.

3. An addition to either study described above could be made in order to determine whether the Making It Strange series teaches metaphoric usage or simply tacitly frees the student to use figurative language in a possibly restrictive composition task. At the conclusion of the experimental period and the posttest, control teachers could use prepared lessons on the values and purposes of figures of speech and reinforce such usage in classroom compositions. A further test could then be administered and the results of the control groups' tests could be compared to the posttests of the experimental group. In this way, it would be possible to separate the role of the Making It Strange series in altering the demand characteristics of composition writing as opposed to augmenting the child's control of figurative usage.

4. In order to determine whether novel figures of speech bring about longer compositions, a study might be conducted in which children complete different composition fragments, some of which use highly metaphoric language and some of which use extremely literal language. The length of the different completions might then be compared in an attempt to determine motivational effects of figurative usage.

5. Since it has been established that children as young as eight years old (Grade 3) do use both frozen and novel figures of speech, it might be interesting to determine when children first begin to use such language. Oral story-telling would represent a feasible approach. Nursery schools, day-care centers, and kindergarten to Grade 2 would provide sufficient children for this type of study.

A further question might be answered in a study such as this: When (at what age) do children use the greatest number of novel figures of speech? Since on a composition task the frequency of novel figures decreases from Grade 3 on, would the peak of usage be at Grade 3 or Grade 2 or at some earlier level? In addition, the pretest composition task might also be extended to Junior High School and Senior High School students as well.

6. A final study suggested by the procedures and materials used in the present experiment would involve an evaluation of the cumulative effects of the Making It Strange series over a four year period of time. At the conclusion of this experience, it would be appropriate to determine if children exposed to this series were able both to use figurative language and to solve open-ended problems more effectively than comparable control children.

Studies such as these suggested above would have to be completed before any more definite conclusions could be reached on either the development of figurative language or the value of the Making It Strange series, and before definite suggestions could be given to classroom teachers on what to do to promote the use of figurative language -- and by extension creativity -- in the children in the elementary schools.

BIBLIOGRAPHY

Armstrong, D.T. "Literary Allusions." Engl. J., 34:218-19, April, 1945.

Asch, Solomon E. "The Metaphor: A Psychological Inquiry" in Person Perception and Interpersonal Behavior, ed. by Tagiuri, R. and Petrullo, L. Stanford: Stanford University Press, Inc., 1958.

Asch, Solomon E. and Nerlove, Harriet. "The Development of Double Function Terms in Children" in Perspectives in Psychological Theory, ed. by Kaplan, B. and Wapner, S. New York: International Universities Press, Inc., 1960, pp. 47-60.

Barlow, J.M., Kerlin, J.R., and Pollio, H.R. Training manual for identifying figurative language. Technical Report No. 1, September, 1971, University of Tennessee, Knoxville.

Bolinger, D. "The Atomization of Meaning," 1965, in Readings in the Psychology of Language, ed. by Jakobovits, L.A. and Miron, M.S. Englewood Cliffs, New Jersey: Prentice-Hall, 1967, pp. 432-448.

Brown, James I. "Vocabulary - Key to Communication." Education, 80:80-84 October, 1959.

Brown, Roger. "Appendix" in Bruner, J.A., Goodnow, J.J. and Austin, G.A. A Study of Thinking. New York: Wiley, 1956.

Brown, Roger. "How Shall a Thing be Called?" Psychol. Rev., 65:14-21, 1958a.

Brown, Roger. Words and Things. Glencoe, Illinois: The Free Press, 1958b.

Brown, Roger. Social Psychology. New York: The Free Press, 1965.

Burkland, C.E. "The Presentation of Figurative Language." Q. J. Speech, 41: 383-390, February, 1964.

Cooper, David. "Concepts from Semantics as Avenues to Reading Improvement." Engl. J., 53:85-90, February, 1954.

Corbett, Edward P.J. Classical Rhetoric for the Modern Student. New York: Oxford University Press, 1965.

Ekstein, Rudolf. Children of Time and Space of Action and Impulse. New York: Appleton-Century Crofts, 1966.

Ferguson, Charles W. "Words as Pictures." Natl. Parent Teachers, 52:14-16

Fine, Harold J., Pollio, H.R., and Simpkinson, C.H. "Figurative language, metaphor and Psychotherapy." Psychotherapy: Theory, Research and Practice. 1973, 10, #1, pp. 87-91.

Gordon, William J.J. Synectics. New York: Harper and Row, 1961.

Gordon, William J.J. "The Metaphorical Way of Knowing," in Education of Vision, ed. by Kepes, G. New York: George Braziller. 1965, pp. 96-103.

Holstein, B.I. Use of metaphor to induce innovative thinking in fourth grade children. Education, 1972, 93, 56-60. (a)

Holstein, B.I. A case study analysis of several children's responses to metaphorical experiences. J. of Educ., 1972, 154, 45-69. (b)

Hughes, T. "Linguistic Approaches to Figurative Speech." Instr. 77:58-59, December, 1967.

Katz, J.J. and Fodor, J.A. "The Structure of a Semantic Theory." Language, 39:170-210, 1963.

Koen, F. "An Intra-Verbal Explication of the Nature of Metaphor." J.V.L.V.B., 4:129-133, April, 1965.

Leedy, J. J. Poetry Therapy. Philadelphia: J.B. Lippincott, 1969.

Lenrow, P.B. "The Uses of Metaphor in Facilitating Constructive Behavior Change." Psychotherapy, 3:145-148, 1966.

Meade, R.A., Haugh, O.M. and Sonkie, O.E. Effective English. Boston: Allyn and Bacon, 1961.

Miller, G.A. "The Magical Number Seven, Plus or Minus Two." Psychol. Rev., 63:81-97, 1956.

Mulder, A. "Wrap Your Ideas in Cellophane." Education, 80:85-87, October, 1959.

Myers, H.A. "The Usefulness of Figurative Language." Q. J. Speech, 26:236-243, April, 1940.

Newton, E.S. and Read, J. "Figurative Language: An Achilles Heel in Reading Comprehension." J. of Reading, 8:65-74, October, 1964.

Roberts, P., Ross, C.S. and Boyd, J. Roberts English Series. New York: Harcourt, Brace and World, 1970.

Sartin, H.W. English Is Our Language. Boston: D.C. Heath and Co., 1968.

Schon, Donald A. Displacement of Concepts. London: Tavistock Publications, 1963.

Sewell, Elizabeth. The Human Metaphor. South Bend, Indiana: University of Notre Dame Press, 1964.

Stelzer, H.G. "Analysis by Metaphor." Q. J. Speech, 51:52-61, February, 1965.

Stern, G. Meaning and Change of Meaning. First published, 1931. Republished Bloomington, Indiana: Indiana University Press, 1954.

Stewig, J.W. "Metaphor and Children's Writing." Elem. Engl. 43:121-34, February, 1966.

Synectics, Inc. Making It Strange. New York: Harper and Row, 1968.

Ullmann, S. "Semantic Universals." In Universals of Language. ed. by Greenberg, J.H. Cambridge, Massachusetts: The M.I.T. Press, 1963.

Upton, A. "Reading and Intelligence." Claremont Reading Conf. Yrbk., 28:137-143, 1964a.

Upton, A. "A Linguistic Approach to Problems of Analysis and Interpretation." J. Comm., 14:118-22, June, 1964b.

Webster, N. Webster's Universal Dictionary of the English Language. New York: World Syndicate Publishing Co., 1937.

Weinreich, U. "Explorations in Semantic Theory" in Current Trends in Linguistics, Vol. III: Theoretical Foundations, ed. by Sebeok, T.A. The Hague: Mouton and Co., 1966.

Wiersma, Mildred. "Ceiling Unlimited." Education, 80:76-79, October, 1959.

Appendix A - Prototype of Composition Tasks Used in Present Study

Name_____ Grade_____

Teacher_____ School_____

Below are five topics for a story. Some suggestions for what to put in the story are included. Choose one topic and write a composition about it. Try to make your story at least a page long.

1. What would you do if all trees disappeared?
   (Describe what it would look like. Think of what trees do for us. What would happen?)

2. Write the adventures of a lazy boy lost in a forest.
   (How would he feel? What would he think about? What would he do? Describe what the forest would look like to him after he realized he was lost.)

3. How would you feel if you were in a pet store and one of the goldfish started to talk to you?
   (What do you think he might say? How would you answer him? Why do you think he talked? Describe the fish.)

4. Write a story called "My Adventures in Space."
   (How do you think you would feel? What would you do? Why? What would you think about? How would your family and friends feel? What would they think about? Describe how the earth looks from your space capsule.)

5. Write a story called "The Coming of Winter."
   (How do you know that winter is near? What can you see? What can you smell? What can you taste? How do you feel? What do you do?)

Appendix B - Additional Work Supported by This Grant

In addition to the work dealing specifically with figurative language in school children, other studies dealing with the role of figurative language in different contexts -- primarily psychotherapy interviews -- were also completed with partial support of this grant. This appendix presents some of these collateral papers, in addition to the Training Manual, which forms the basis of the experimental program.